

# Classificação de Morangos a partir de Imagens usando Classificadores Supervisionados e Redes Convolucionais como Extratoras de Características

Érika Kayoko Hamaguti

Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista (UNESP)  
Rio Claro-SP, Brasil  
erika.k.hamaguti@unesp.br

Fabrizio Aparecido Breve

Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista (UNESP)  
Rio Claro-SP, Brasil  
fabrizio.breve@unesp.br

**Resumo**—As técnicas tradicionais utilizadas para identificar a maturidade do morango são baseadas na aparência ou na composição química. Porém, essas técnicas são demoradas e caras. A classificação automática da maturidade do morango pode ajudar os agricultores a colher no tempo ideal, melhorar a produtividade e ter um melhor gerenciamento do negócio. Este trabalho tem como objetivo realizar vários testes em modelos de CNNs como extratores de características e modelos de aprendizado supervisionado para encontrar os modelos mais adequados para a tarefa de classificação de morangos maduros.

**Área:** *Inteligência Computacional.*

## I. INTRODUÇÃO

O consumo de alimentos saudáveis e de forma equilibrada é essencial para ter uma boa saúde [1]. Com o tempo, os consumidores brasileiros vêm valorizando cada vez mais os produtos relacionados à qualidade de vida [2]. Um dos alimentos preferidos pelos consumidores é o morango, por conta de suas características como aroma, textura suculenta e doçura. A maturação do morango tem sido avaliada, convencionalmente, por especialistas ou pesquisadores e são verificados aspectos como cor, textura e composição química para determinar a qualidade da fruta. Apesar de ter uma precisão satisfatória, esses métodos geralmente são destrutivos, demorados e trabalhosos. Por conta disso, é preciso criar um método rápido e não destrutivo para avaliar a maturação de morangos [3]. Para esse problema é possível utilizar algoritmos de aprendizado de máquina que possam identificar automaticamente os morangos prontos para serem colhidos.

Gao *et al.* [3] utilizaram *hyperspectral imaging* (HSI) para extrair informações das imagens de morango e *Support Vector Machines* (SVM) para classificação de maturidade, e foi obtido o ROC maior do que 95%. Behera, Rath e Sethy [4] combinam os recursos LBP, HOG e GLCM com os classificadores *K-Nearest Neighbors* (KNN), SVM e *Naïve Bayes* para comparar os resultados obtidos na classificação de maturidade de mamão, e o melhor resultado encontrado foi o KNN e HOC com 100% de precisão e com tempo de treinamento de 0,0995 segundos. Benmouna *et al.* [5] utilizam Visible/Near Infrared (Vis/NIR) para extrair características das imagens e utilizam

redes neurais artificiais (RNA), SVM e KNN para classificar a maturidade da maçã Fuji, e os resultados foram 89.5%, 95.93%, e 91.68% de *correct classification rate* (CCR) para RNA, SVM, e KNN, respectivamente.

Este trabalho tem como objetivo encontrar os modelos mais adequados dentro de CNNs como extratores de características e de modelos de aprendizado supervisionado para o problema de identificação de morangos maduros utilizando imagens digitalizadas. Este tema foi escolhido, pois a primeira autora trabalhou em sua graduação com uma estufa automatizada portátil, e poderia incluir na mesma a tarefa de verificação periódica de maturidade. Por conta de ser uma estufa pequena, foi escolhido o morango para ser trabalhado.

## II. CONCEITOS E TÉCNICAS

### A. Aprendizado Supervisionado

O aprendizado supervisionado é uma categoria de *Machine Learning* que utiliza conjuntos de dados que são rotulados com base no conhecimento de especialistas de uma determinada área [6]. Esse aprendizado infere uma função ou regra para poder atribuir uma classe a um objeto nunca visto antes a partir de dados de treinamento previamente classificados [7,8]. Ou seja, o objetivo do aprendizado supervisionado é prever uma saída para cada dado de entrada, baseando-se no padrão aprendido em cada uma das classes no treinamento realizado com dados rotulados [8].

### B. Redes Neurais Convolucionais

As redes neurais convolucionais, do inglês *Convolutional Neural Network* (CNN) [9], são baseadas no córtex visual do cérebro e têm sido muito eficientes no reconhecimento de imagens, incentivando o seu uso em várias áreas de pesquisa. Uma CNN é formada basicamente por camadas convolucionais, *pooling* e camadas totalmente conectadas para extrair as características dos dados de forma automática [10].

## III. METODOLOGIA DE DESENVOLVIMENTO

Neste trabalho, o conjunto de imagens utilizado foi o Strawberry-DS, que possui 247 imagens em RGB com

resolução de 3840 x 2160 pixels [11]. As imagens foram separadas e rotuladas como “Colher” e “Nao\_colher”, sendo “Colher” as imagens que têm pelo menos um morango para colher, e “Nao\_colher” as imagens sem morangos prontos para colher. No total, o número de imagens de “Colher” são 170 e de “Nao\_colher” são 77.

Na extração de características, foram utilizadas várias CNNs, como VGG16, VGG19 [12], ResNet50, ResNet101 e ResNet152 [13]. Elas são usadas sem a camada de classificação, com pesos pré-treinados no conjunto de dados ImageNet [14].

Os modelos de aprendizagem supervisionado utilizados foram *Linear Support Vector Machines* (LSVM), *Logistic Regression* (LR), *K-Nearest Neighbors* (KNN), *Naïve Bayes* (NB), *Decision Tree* (DT) e *Random Forest* (RF). Para o treinamento desses modelos, foi utilizado o *Cross Validation* com 5 *folds*. Primeiro foram realizados treinamentos dos modelos com dados desbalanceados, e depois com dados balanceados. Para o balanceamento de dados, foi utilizada a técnica de aumento de dados, com espelhamento na horizontal, rotação e zoom, dentro da classe “Nao\_colher” e foi aumentado até o número de imagens equivaler ao número da classe “Colher”. As métricas de comparação escolhidas para mostrar neste artigo foram a média da acurácia e a média da F1-Score.

A plataforma Google Colab foi utilizada na programação dos modelos. Ela utiliza Python versão 3 como linguagem de programação e executa os códigos em nuvem do Google.

#### IV. RESULTADOS

Os resultados obtidos até o momento são exibidos na Tabela 1.

TABELA 1 – RESULTADOS OBTIDOS.

Técnicas Aplicadas	Sem balanceamento		Com balanceamento	
	Acurácia	F1-Score	Acurácia	F1-Score
VGG16 + LSVM	0.87347	0.91257	0.87347	0.91206
VGG16 + LR	0.85714	0.90276	0.86939	0.91146
VGG16 + KNN	0.63673	0.69073	0.63673	0.69073
VGG16 + NB	0.74694	0.84617	0.74694	0.84617
VGG16 + DT	0.62448	0.72958	0.67347	0.76632
VGG16 + RF	0.58367	0.69275	0.62857	0.73189
VGG19 + LSVM	0.84898	0.89498	0.84898	0.89559
VGG19 + LR	0.86531	0.90860	0.85306	0.90012
VGG19 + KNN	0.70204	0.78474	0.70204	0.78474
VGG19 + NB	0.73061	0.83761	0.73469	0.83982
VGG19 + DT	0.64490	0.73652	0.62857	0.72819
VGG19 + RF	0.66939	0.75334	0.66122	0.75224
ResNet50 + LSVM	0.65714	0.69892	0.65306	0.71420
ResNet50 + LR	0.68980	0.77961	0.71837	0.80071
ResNet50 + KNN	0.69796	0.80623	0.69796	0.80623
ResNet50 + NB	0.65306	0.75470	0.64898	0.75251
ResNet50 + DT	0.64490	0.73502	0.63673	0.73184
ResNet50 + RF	0.64898	0.74106	0.64490	0.73411
ResNet101 + LSVM	0.66531	0.76999	0.65306	0.74391
ResNet101 + LR	0.68571	0.77610	0.69796	0.78546
ResNet101 + KNN	0.68571	0.80323	0.68571	0.80323
ResNet101 + NB	0.66531	0.77137	0.67347	0.77823
ResNet101 + DT	0.61224	0.70793	0.6	0.71469
ResNet101 + RF	0.62857	0.73221	0.55918	0.69011
ResNet152 + LSVM	0.65306	0.72632	0.71020	0.80600

Técnicas Aplicadas	Sem balanceamento		Com balanceamento	
	Acurácia	F1-Score	Acurácia	F1-Score
ResNet152 + LR	0.66531	0.76751	0.64490	0.75175
ResNet152 + KNN	0.69388	0.80785	0.68980	0.80472
ResNet152 + NB	0.64490	0.74767	0.62857	0.73957
ResNet152 + DT	0.6	0.70175	0.58367	0.68542
ResNet152 + RF	0.59184	0.69990	0.60408	0.70129

Com base nos testes realizados no conjunto Strawberry-DS, a melhor combinação encontrada foi LinearSVM com o extrator VGG16. Um ponto interessante é que essa combinação obteve o valor da F1-Score ligeiramente melhor quando não tem balanceamento de dados, com 91,26% de média.

#### V. CONSIDERAÇÕES FINAIS

Os próximos passos deste trabalho são testar com outro conjunto de imagens, chamado StrawDI\_DB1, e realizar vários testes com outros CNNs como extratores de características e modelos de *Machine Learning* na classificação para que seja possível encontrar os modelos mais adequados para esta tarefa.

#### REFERÊNCIAS

- [1] BEVILACQUA, H. E. C. R. Classificação das hortaliças. In: CASTANHEIRO, ALM; BEVILACQUA, HECR; SHIRAKI, JN (Coords.). **Horta: cultivo de hortaliças**. São Paulo: Prefeitura do Município de São Paulo, Secretaria Municipal do Verde e do Meio Ambiente, p. 1039-1042, 2006.
- [2] VIEIRA, D. F. A. **Catálogo brasileiro de hortaliças**: saiba como plantar e aproveitar 50 das espécies mais comercializadas no país. Brasília: Embrapa Hortaliças: SEBRAE, 2010. p. 59.
- [3] GAO, Z. *et al.* Real-time hyperspectral imaging for the in-field estimation of strawberry ripeness with deep learning. **Artificial Intelligence in Agriculture**, v. 4, p. 31-38, 2020.
- [4] BEHERA, S. K.; RATH, A. K.; SETHY, P. K. Maturity status classification of papaya fruits based on machine learning and transfer learning approach. **Information Processing in Agriculture**, v. 8, n. 2, p. 244-250, 2021. DOI:10.1016/j.inpa.2020.05.003.
- [5] BENMOUNA, B. *et al.* Convolutional neural networks for estimating the ripening state of Fuji apples using visible and near-infrared spectroscopy. **Food and Bioprocess Technology**, v. 15, n. 10, p. 2226-2236, 2022. DOI:10.1007/s11947-022-02880-7.
- [6] HAYKIN, S. **Redes Neurais**: Princípio e Prática. Bookman, 2001.
- [7] ALPAYDIN, E. **Introduction to machine learning**. MIT Press, 2004.
- [8] MITCHELL, T. **Machine learning**. McGraw Hill, 1997.
- [9] LECUN, Y. *et al.* Object Recognition with Gradient-Based Learning. In: Shape, Contour and Grouping in Computer Vision. **Lecture Notes in Computer Science**, 1999. v. 1681. Springer, Berlin, Heidelberg.
- [10] ALZUBAIDI, L. *et al.* Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. **Journal of big Data**, v. 8, n. 1, p. 1-74, 2021.
- [11] EL-BENDARY, N.; ELHARIRI, E. **Strawberry-DS**, Mendeley Data, V1, DOI: 10.17632/z6dtdfpz8.1, 2022.
- [12] SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **Computational and Biological Learning Society**, p. 1-14, 2015.
- [13] HE, K. *et al.* Deep residual learning for image recognition. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. 2016. p. 770-778.
- [14] RUSSAKOVSKY, O. *et al.* ImageNet Large Scale Visual Recognition Challenge. **International Journal of Computer Vision (IJCV)**, v. 115, n. 3, p. 211-252, 2015.